ELSEVIER

Immunotechnology 4 (1999) 279–285

IMMUNOTECHNOLOGY

Commentary

Complementarity-determining region (CDR) implantation: a theme of recombination

Eskil Söderlind [a,*], Mats Ohlin [b], Roland Carlsson [a]

[a] *BioInvent Therapeutic AB, SE-223 70 Lund, Sweden*
[b] *Department of Immunotechnology, Lund University, SE-220 07 Lund, Sweden*

Received 30 November 1998; accepted 1 December 1998

Abstract

A novel technology in the area of antibody engineering has been developed which allows for the creation of new types of antibody molecules. It is called complementarity-determining region (CDR) implantation and permits the random combination of CDR sequences formed in vivo into a single master framework. Thus, totally new gene combinations can be produced and used in selection processes. The result is a genetic variability which is extremely large, even exceeding the natural variability found in the immune system. In this commentary, CDR implantation is presented and the technology is discussed. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Antibody engineering; CDR implantation; Complementarity-determining region; Genetic variation

1. Introduction

A central problem in the area of antibody engineering is the creation of genetic variation. In the traditional setting of the technology, gene sequences are harvested from the immune system where the joining of distinct VJ and VDJ segments build up the genes for variable light (V_L) and variable heavy (V_H) chains. The in vivo genetic variation is thus based on the process of recombination, a process which allows for large variability to be formed from a limited number of building blocks. This naturally occurring genetic variability can be used as a source for in vitro selection of highly specific antibody fragments. Furthermore, such selections have been revolutionised by the introduction of phage display [1–5] through which a repertoire of antibody

* Corresponding author. Tel.: +46-46-2868577; fax: +46-46-2868570.
E-mail address: es@bioinvent.com (E. Söderlind)

1380-2933/99/$ - see front matter
PII: S1380-2933(98)00027-X

fragments can be displayed on the surface of filamentous phage, linking the binding phenotype to its corresponding genotype. This provides the possibility to directly clone, analyse and modify the gene encoding the specificity.

Successful selection and cloning of antibody fragments with a desired specificity depends on the presence of the corresponding genes encoding this specificity in the library. The use of immunised genetic libraries (i.e. libraries derived from individuals with an immune response to a particular antigen) improves the probability of successful selections since this immune repertoire has an increased frequency of clones reacting with the antigen. Of course, the source of this type of genetic library, based on human antibody genes, is limited for ethical and practical reasons. Another route for successful selection of antibody fragments is to construct, in vitro, libraries with extended genetic variation. These libraries contain large genetic variability and their use increases the probability to find specificities against any antigen, including specificities not found in an immunised library (e.g. against self antigens).

The in vitro constructed genetic variation in antibody gene libraries can be targeted to the complementarity-determining regions (CDR) of the immunoglobulin gene [6–9], which encode loops building up the combining site of the antibody molecule. Previous technologies used for this type of targeted genetic variation are based on the use of degenerate synthetic oligonucleotides (oligos) which are of non-natural origin. Although careful construction of these oligos often is the case, their encoded variation does not resemble all the variation found in naturally occurring immunoglobulin genes. The reason for this is that our knowledge of the immune repertoire is not complete and any construction will be biased and influenced by our present, limited understanding of the genetic variation in the immune system. Thus, there is a possibility of a non-optimal construction of synthetic oligos. In contrast, the natural genetic variation formed in vivo is already optimised with regard to functionality since it has been processed in the immune system. The use of such in vivo genetic variation is preferred, but no technology has so far been developed which allows for in vivo CDR segments to be combined randomly. Here we present such a technology.

2. CDR implantation—a novel technology for gene libraries

We have developed a unique system called CDR implantation in which natural gene segments are randomly combined into a defined master framework [10] (Söderlind et al., in preparation), selected for optimal expression in bacterial and phage systems. The technology includes the synthesis of genes encoding antibody variable domains by overlap extension polymerase chain reaction (PCR) of a set of single-stranded oligonucleotides. Some of the primers are amplified from natural immunoglobulin genes and contain CDR sequences (see Fig. 1 for details).

The CDR-implantation system excludes the need for oligo construction and synthesis for in vitro genetic diversity, but uses genes formed in vivo, harvested from the immune system. Some of these genes have undergone affinity maturation processes and carry mutations at different frequencies, whereas some others have germline configuration. In addition, further variability can be found since the immune system is capable of modifying the genes by insertion mutagenesis, but only a small minority of the repertoire will carry evidence of such events [11–13]. In contrast to present systems, where genes formed in vivo encoding the V_H and V_L chains are combined in a dimeric format [14,15], the CDR-implantation system allows for combination of in vivo formed gene segments in a hexameric format for the creation of gene diversity (Fig. 2), with a dramatic increase in variation as the result.

CDR implantation is based on previous work [10,16,17] on overlap extension PCR as the method for gene synthesis and to randomly combine different in vivo formed CDR sequences (Fig. 3). In the overlap extension gene synthesis process, it is possible to use some oligos carrying degenerate regions [16]. This points to the important fact that oligos with genetic variation are not detrimental for the extension process and full-

length gene sequences can be obtained with fully variegated sequences at defined positions. A central problem in the use of overlapping PCR for gene library construction is the design of the participating single-stranded oligos [17]. By design means the determination of the length of each individual oligo participating in the overlap extension process and not the previously mentioned construction of in vitro genetic variation. A single gene can be assembled by overlap extension, and selection processes can be applied to identify clones with an open reading frame (ORF). However, if a genetic library is to be created which is to encode several genetic variants, each with an ORF, the product from such an overlap extension process needs to contain a minimum of mutations/deletions. To ensure such a high frequency of ORFs, careful oligo design is a prerequisite. In such design, parameters such as homo- and heterodimerisation as well as mispriming have to be analysed. Furthermore, if the oligo design allows for unwanted hybridisation stability to be 10–25% of the wanted oligo hybridisation, there is a good probability for successful gene synthesis [17]. The designed primers described in [17] were used in the CDR-implantation library construction and resulted in 75% genes with ORFs (Söderlind et al., in preparation). Furthermore, the quality of the oligos in the overlap PCR process has to be good (i.e. they should contain a minimum of deletions/mutations). We have observed that if, instead of using synthetic DNA in the overlap PCR process, DNA is produced by PCR (which is the case when CDR regions are amplified in the CDR-implantation process), the frequency of genes with an intact ORF increases [10] (Kobayashi and Söderlind, unpublished).

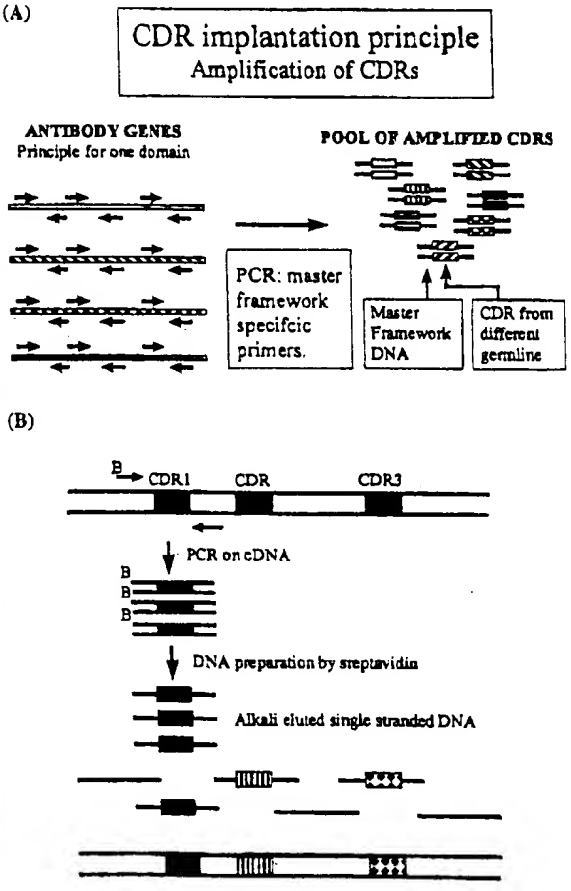

Fig. 1.

Fig. 1. Principle layout for the CDR-implantation technology. DNA sequences encoding CDR loops can be amplified by the use of primers (A) which are specific for a master framework. The resulting pool of amplified CDR sequences can be used in a subsequent overlap extension PCR process (B) where they randomly combine with other CDR sequences of different germline origin. The DP-47 and DPL-3 germline genes were selected to encode the master framework. To incorporate the CDRs into these master frameworks, oligonucleotides based on the DP-47 and DPL-3 germline genes were used in a PCR. For each CDR amplification, an oligonucleotide pair was designed to amplify the CDR as well as to allow for one strand of the PCR product to be used in gene assembly. One of the primers used was biotinylated at the 5′ end. A human cDNA library derived from peripheral blood lymphocytes and tonsils was used as template for the CDR sequences. Single-stranded DNA was initially prepared in order to be able to utilise the PCR-amplified CDR in an assembly reaction using overlapping oligonucleotides. This was performed by affinity chromatography on the biotinylated strand; the biotinylated PCR product was affinity captured on the streptavidin-coated matrix and the column was washed to remove any remaining template DNA. The non-biotinylated strand was finally eluted by denaturing the DNA with alkali. The eluted single-stranded DNA was subsequently used in the assembly reaction. Five overlapping internal oligonucleotides and two amplification primers were used in the assembly PCR [16]. The three internal oligonucleotides encoding the CDRs were prepared as described above, and used together with the other oligonucleotides in an overlap extension process.

3. High functionality with gene segments of natural origin

Compared with gene segments formed in vivo, constructed synthetic oligos suffer from several limitations:

1. A synthetic, fully degenerate oligo will encode all three stop codons; to overcome this problem, limited variation can be constructed at the third nucleotide in the codon [6,7]. All amino acids will be present in such a synthetic library but not allow all codons to be used, biasing the usage of codons. Gene segments of natural origin do not suffer from these limitations.
2. If codon-based mutagenesis approaches are used [18,19], the synthetic variation still needs to be constructed and will not fully resemble all the variation found in genes formed in vivo.
3. When constructed synthetic degenerate oligos are used, there is a possibility that genes encoding antigenic epitopes are produced. The use of in vivo gene segments will probably reduce this possibility, since polypeptides encoded by these segments have been adapted and processed by the immune system.

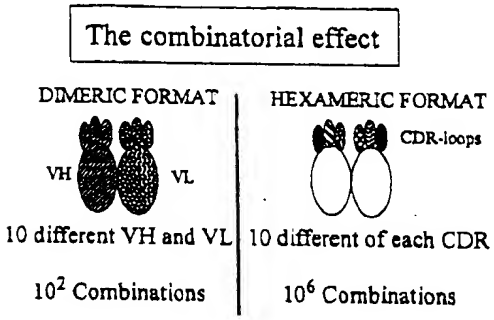

Fig. 2. The combinatorial effect. By random combination of discrete gene segments, encoding specific peptide stretches, it is possible to generate extended variation. In the immune system, the variable heavy (V_H) and variable light (V_L) domains are combined in a dimeric format, with an increased variation as the result. In the CDR-implantation technology, individual CDR sequences can be randomly combined in a hexameric format, with a further creation of variability as the result.

In vivo formed DNA sequences are a good source for the creation of molecules with new function and the concept of random recombination of in vivo formed gene segments was also recently described in the DNA-shuffling technology [20]. Here, a family of four genes encoding moxalactamase could be recombined into a new gene containing segments from the different genes and this chimeric gene encoded a moxalactamase with improved function. CDR implantation allows for random combination of defined gene segments and differs from DNA shuffling in this respect. Nevertheless, in both cases in vivo formed gene segments can be recombined and a larger part of sequence space can be exploited.

4. Structural considerations

In the germline repertoire certain combinations of CDR1 and CDR2 are inherited as units as this part of the human immunoglobulin gene does not involve recombination or gene conversion events. The sequence and length of each such hypervariable region determines its basic fold, or so-called canonical structure [21]. Since these CDRs are inherited pair-wise, only certain combinations of hypervariable regions will make up the original antibody repertoire in humans. Consequently, certain combinations of canonical structures seem to dominate the immune repertoire [22]. For instance, antibodies derived from the very abundant V_H3 gene family, usually employ canonical structures of type 1 and 3 for CDR1 and CDR2, respectively. In the context of CDR implantation, it is possible to maintain the original canonical loop structure associated with the chosen framework or to introduce novel loop folds. As previously demonstrated, it is possible through careful primer design to specifically amplify V_H CDR originating from specific immunoglobulin gene families [10]. We have in this way incorporated V_H CDR1 and V_H CDR2 derived solely from genes belonging to the V_H3 gene family into framework derived from the 3 to 23 germline gene loci [10] and by using this principle also into framework derived from the 3 to 30 germline gene loci [23]. For example, as shown in Fig. 4, it is

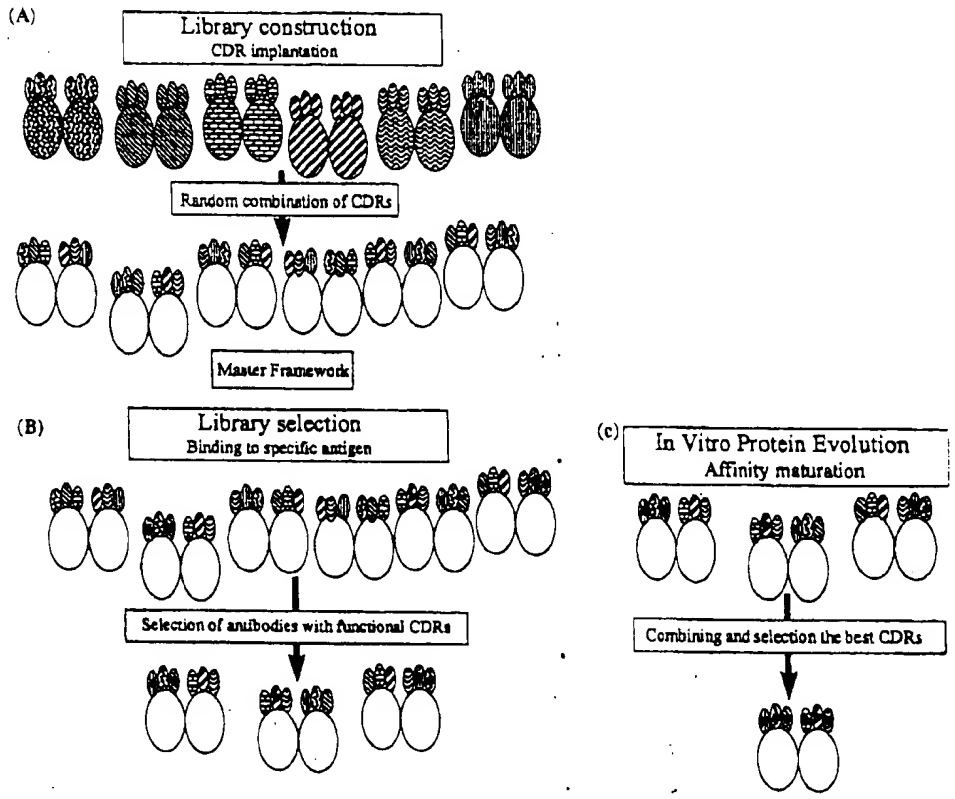

Fig. 3. The recombination process using the CDR-implantation technology. (A) From a set of natural antibodies it is possible to randomly combine CDR sequences from these to a defined single framework region. CDR sequences from position 1 in, for example, the V_L domain is always placed in position 1 in the master framework V_L domain, etc. The result is an antibody gene library. (B) Selection of specific antibody fragments from the CDR-implantation library is carried out and CDR loops that are participating in the antigen binding are marked with a black dot. (C) By the use of the CDR implantation it will be possible to combine the best CDR loops from different antibody fragments into one single molecule in which all six CDR loops participate in antigen binding. Here, a new library is produced and the antibody fragments with improved characteristics (i.e. affinity and specificity) are selected.

obvious that the majority of CDR2 originating from human V_H3 sequences may be amplified, while those originating from other families will not (except in the few cases where the somatic mutation process has modified critical bases in the template close to the 3′ end of the primers). In addition, we have noticed some length variability of lambda light chain loops in certain specific clones selected, from CDR-shuffled antibody libraries (Söderlind et al., in preparation), causing us to believe that such variability is indeed compatible with a functional immunoglobulin fold. Furthermore, the introduction of divergent canonical structures into a master framework would provide insight into the possibility to create new types of functional variability using the basic antibody scaffold. Indeed, it is possible to incorporate variability derived from other germline gene families and thus different canonical loop structures into a specific antibody gene by modification of the primer design (Ohlin et al., in preparation). This approach might also be considered as an analytical tool which can be used for the study of CDR loops in relation to framework structures.

codon	36	37	38	39	40	41	42
VH CDR2 primers		GTC	CGC	CAG	GCT	CCA	--->
VH1 consensus	TGG	..G	..A	...	..Y	..Y	GGA
VH2 consensus	...	A..	..T	...	C.C	...	..R
VH3 consensus	...	...	...	..R	...	...	..S
VH4 consensus	...	A..	..S	...	CMC	...	..G
VH5 consensus	...	..G	...	...	ATG	..C	..G
VH6 consensus	...	A..	A.G	...	T.C	...	TCG
VH7 consensus	...	..G	..A	...	..C	..T	...

CDR2

66	67	68	69	70	71	72	73	74	75
<---	TTC	ACC	ATC	TCC	AGA	GAC	AAT	TCC	
AGA	G..	..S	..K	A..	RSG	...	.NR	..Y	AYR
..G	C..	...	...	W..	.AG	...	.CC	...	.AA
M..	...	...	...	..M	...	..Y	R.Y	K.M	.A.
C..	G.Y	...	..A	..A	GT.	...	.CG	..Y	.AG
CAS	G..	...	...	..A	GCY	...	..G	...	.TC
C..	A.A	...	...	AA.	CC.	...	.CA	...	.AG
C.G	..T	GT.	T..	...	TTG	...	.CC	..T	GTC

Fig. 4. Primers used for amplification of V_H CDR2 for incorporation in the CDR-implantation library. By using these primers, design codons 42–66 (including CDR2) can be amplified from naturally occurring immunoglobulin-encoding genes. The primer design is such that only genes belonging to germline gene family V_H3 would be expected to be efficiently amplified due to mismatches at the 3′ end of the other gene families. Some somatic mutation of the template would be expected to allow amplification of the CDR belonging to gene families other than V_H3. The primer sequences are compared to consensus sequences (here represented by the bases found at > 90% of loci belonging to each V_H gene family) of functional germline genes, as described in the VBASE directory [25]. Residue numbering according to Chothia et al. [21] is shown at the top of the figure. Nucleotides are given by standard notation (Y = C,T; R = A,G; S = G,C; M = A,C; W = A,T; K = G,T; N = A,G,C,T).

5. In vitro optimisation

The process of CDR implantation involves random recombination of CDR sequences without the use of any restriction enzymes. The CDR sequences are amplified from natural immunoglobulin genes in a PCR process and linked to the master framework by a PCR overlap extension process. In order to optimise the antibody fragment with regard to specificity and affinity, it is possible to repeat this process of CDR implantation from a pool of selected antibody fragments. Thus, in such a way, already selected CDR sequences will be recombined and this has the potential to create molecules with optimised combination of CDR sequences (Fig. 3)—i.e. the sequences contributing to affinity and specificity could be collected in one single molecule. Such molecules would exhibit excellent specificity and affinity. Several rounds of repeated CDR implantation might be necessary. It is important to realise that there is no need for any sequence information concerning the participating CDR sequences when performing the CDR-implantation process.

The CDR-walking technology [24] has been tested for the optimisation of antibody fragments. Here, in a first step, a pool of best CDRs are selected and in the next step a second CDR is optimised and so forth. This strategy for sequential combination of optimised CDR sequences resulted in antibody fragments with improved affinities [24]. The CDR-walking and the CDR-implantation technology differ; the sequences used in the CDR-implantation technology are natural gene sequences formed in vivo, whereas the CDR-walking system depends on constructed degenerate oligonucleotides for the creation of genetic variability. Furthermore, in CDR implantation there is a possibility to simultaneously combine between two and six CDR sequences and to select the best combinations of CDRs. In the CDR-walking alternative, CDR sequences are fixed and the procedure moves to the optimisation of the next CDR. Thus, there is a first choice of individual CDR sequences which form the base for further optimisation. However, in theory, these chosen CDR regions might not contain optimal sequences in combination with one or several other CDR sequences. Thus, it is important to utilise a set of different CDR sequence combinations, in a library format, to select the best combinations.

6. Concluding remarks

When developing the CDR-implantation concept, a central question was whether or not the increased variability introduced through recombination of all six CDRs would result in a functional variability—i.e. would the library constructed allow for specific selection of antibody fragments against a variety of antigens. In-

deed, initial analysis of a large antibody fragment library, created using the CDR-implantation technology, demonstrates that high-affinity antibodies against different types of antigens can be selected (Söderlind et al., in preparation). Thus, the different topologies in the antibody-combining site build a variety of surfaces which allow for different antigens to be bound.

When comparing variability between antibodies generated from the CDR-implantation technology with conventional antibodies, CDR1 and CDR2 will be found in combinations not present in natural antibodies (Söderlind et al., in preparation). The base and amino acid sequences in the CDR regions also include mutations generated through in vivo somatic mutations, which adds a further level of complexity and increases variability. The potential variability that can be generated by this type of recombination is enormous. If we would recombine, for example, 1000 different CDRs in each position, the possible variability would be 10^{18}, by far exceeding what can be salvaged and handled in any antibody gene library. Thus, the potential variability that can be created using the CDR-implantation technology is essentially unlimited. This means that members in a library created by this technology have a great probability to be different, resulting in a library with a high degree of variability. Furthermore, since this variability is presented in a scaffold that expresses and folds well in *Escherichia coli*, the variability at the functional protein level is also expected to be good. Thus, the CDR-implantation technology has a large potential for antibody development and engineering.

References

[1] Smith GP. Science 1985;228:1315–7.
[2] Parmley SF, Smith GP. Gene 1988;73:305–18.
[3] McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Nature 1990;348:552–4.
[4] Barbas, III CF, Kang AS, Lerner RA, Benkovic SJ. Proc Natl Acad Sci USA 1991;88:7978–82.
[5] Söderlind E, Lagerkvist AC, Duenas M, Malmborg AC, Ayala M, Danielsson L, Borrebaeck CAK. Biotechnology 1993;11:503–7.
[6] Barbas, III CF, Bain JD, Hoekstra DM, Lerner RA. Proc Natl Acad Sci USA 1992;89:4457–61.
[7] Hoogenboom HR, Winter G. J Mol Biol 1992;227:381–8.
[8] Garrard LJ, Henner DJ. Gene 1993;128:103–9.
[9] Griffiths AD, Williams SC, Hartley O, Tomlinson IM, Waterhouse P, Crosby WL, Kontermann RE, Jones PT, Low NM, Allison TJ, et al. EMBO J 1994;13:3245–60.
[10] Jirholt P, Ohlin M, Borrebaeck CAK, Söderlind E. Gene 1998;215:471–6.
[11] Goossens T, Klein U, Küppers R. Proc Natl Acad Sci USA 1998;95:2463–8.
[12] Ohlin M, Borrebaeck CAK. Mol Immunol 1998;35:233–8.
[13] Wilson PC, de Bouteiller O, Liu YJ, Potter K, Banchereau J, Capra JD, Pascual V. J Exp Med 1998;187:59–70.
[14] Huse WD, Sastry L, Iverson SA, Kang AS, Alting-Mees M, Burton DR, Benkovic SJ, Lerner RA. Science 1989;246:1275–81.
[15] Clackson T, Hoogenboom HR, Griffiths AD, Winter G. Nature 1991;352:624–8.
[16] Söderlind E, Vergeles M, Borrebaeck CAK. Gene 1995;160:269–72.
[17] Kobayashi N, Söderlind E, Borrebaeck CAK. Biotechniques 1997;23:500–3.
[18] Huse WD, Yelton DE, Glaser SM. Int Rev Immunol 1993;10:129–37.
[19] Virnekäs B, Ge L, Pluckthun A, Schneider KC, Wellnhofer G, Moroney SE. Nucleic Acids Res 1994;22:5600–7.
[20] Crameri A, Raillard SA, Bermudez E, Stemmer WP. Nature 1998;391:288–91.
[21] Chothia C, Lesk AM, Gherardi E, Tomlinson IM, Walter G, Marks JD, Llewelyn MB, Winter G. J Mol Biol 1992;227:799–817.
[22] Lara-Ochoa F, Almagro JC, Vargas-Madrazo E, Conrad M. J Mol Evol 1996;43:678–84.
[23] Ohlin M, Jirholt P, Thorsteinsdottir HB, Söderlind E, Borrebaeck CAK. In: Antibody Engineering—New Technology, Application and Commercialisation. London: IBC Business Publishing (in press).
[24] Barbas CF, Hu D, Dunlop N, Sawyer L, Cababa D, Hendry RM, Nara PL, Burton DR. Proc Natl Acad Sci USA 1994;91:3809–13.
[25] Tomlinson IM. VBASE. The database of human antibody genes. URL:http://www.mrc-cpe.cam.ac.uk/imt-doc/public/INTRO.html